

A SPECTRAL ANALYSIS OF SUBSPACE ENCHANCED PRECONDITIONERS

TAO ZHAO*

Abstract. It is well-known that the convergence of Krylov subspace methods to solve linear system depends on the spectrum of the coefficient matrix, moreover, it is widely accepted that for both symmetric and unsymmetric systems Krylov subspace methods will converge fast if the spectrum of the coefficient matrix is clustered. In this paper we investigate the spectrum of the system preconditioned by the deflation, coarse correction and adapted deflation preconditioners. Our analysis shows that the spectrum of the preconditioned system is highly impacted by the angle between the coarse space for the construction of the three preconditioners and the subspace spanned by the eigenvectors associated with the small eigenvalues of the coefficient matrix. Furthermore, we prove that the accuracy of the inverse of projection matrix also impacts the spectrum of the preconditioned system. Numerical experiments emphasized the theoretical analysis.

Key words. spectrum, coarse space, deflation, preconditioner, perturbation analysis, projection matrix, iterative solvers, domain decomposition

1. Introduction. We consider the iterative solution of a linear system

$$Ax = b,$$

where $A \in \mathbb{R}^{n \times n}$ is symmetric and positive definite (SPD). It is well-known that the convergence of Krylov subspace methods for solving linear systems depend on the eigenvalue distribution of A . Recently, several studies [2, 3, 5, 7, 8, 10–14, 19] have shown that by removing the subspace spanned by the eigenvectors corresponding to several small eigenvalues from the Krylov search space makes the spectrum more clustered and consequently the convergence is improved.

In this paper, we refer to matrix of the form

$$(1.1) \quad E = Z^T A Z, \quad Z \in \mathbb{R}^{n \times r}.$$

as a projection matrix, and the subspace spanned by the columns of Z is referred to as a coarse space. In an ideal situation, the coarse space contains the vectors corresponding to the lower part of the spectrum that is responsible for the stagnation of Krylov subspace methods. As shown in this paper, the preconditioned systems will have the desired properties when the preconditioner is enhanced with the ideal coarse subspace. In contrast to the general coarse space, we refer to the coarse space spanned by the eigenvectors associated with several small eigenvalues of A as the exact coarse space.

Next we briefly mention a few existing approaches that take the form of a standard precondition with an additive coarse space enhancement. In [3, 12], the deflation preconditioner is defined by

$$(1.2) \quad P_D = I - A Z E^{-1} Z^T.$$

Obviously $P_D A$ is singular since P_D is singular. Fortunately, Krylov subspace methods converge for singular linear systems as long as they are consistent; furthermore, zero eigenvalues do not impact the convergence since the corresponding eigenvectors never enter the Krylov subspace [12].

*Department of Computer Science, University of Colorado Boulder, CO 80309, USA (tao.zhao@colorado.edu)

Instead of zero out the small eigenspace in the deflation method, the preconditioners based on coarse correction shift the small eigenvalues to the large ones. In [19], the coarse correction preconditioner in domain decomposition method is defined by

$$(1.3) \quad P_C = I + ZE^{-1}Z^T.$$

The abstract additive coarse correction is $M^{-1} + ZE^{-1}Z^T$, where M is the sum of the local solves in each subdomain. Another popular preconditioner in domain decomposition method is the abstract balancing preconditioner [3]

$$(1.4) \quad P_{BNN} = (I - ZE^{-1}Z^T A)M^{-1}(I - AZE^{-1}Z^T) + ZE^{-1}Z^T.$$

The adapted deflation preconditioner [19] is defined as

$$P_{ADEF1} = M^{-1}P_D + ZE^{-1}Z^T.$$

It is shown in [19] that P_{ADEF1} is cheaper than P_{BNN} but is as robust as P_{BNN} . Moreover, $P_{BNN}A$ and $P_{ADEF1}A$ have an identical spectrum. In both P_{BNN} and P_{ADEF1} , the first term corresponds to the fine space and the second term is for the coarse space, therefore they are called two-level preconditioners. Throughout this paper, we restrict our analysis to one-level methods. Let M in P_{ADEF1} be I . We define P_A as

$$(1.5) \quad P_A = I - AZE^{-1}Z^T + ZE^{-1}Z^T.$$

Let X be any basis of a coarse space not Z . Obviously the following identity

$$X(X^TAX)^{-1}X^T = Z(Z^TAZ)^{-1}Z^T$$

holds in exact arithmetic since there exists a nonsingular matrix $C \in \mathbb{R}^{r \times r}$ such that $X = ZC$. This implies that P_D , P_C and P_A are determined uniquely by the coarse space. Thus we can choose an appropriate basis to form Z and then to construct a preconditioner with certain desirable properties.

As we will see, if an approximate coarse space is used to construct preconditioners, then the spectrum of the preconditioned systems is related to the angle between the approximate and exact coarse spaces. We also prove that the coarse correction and adapted deflation preconditioners are more robust than the deflation preconditioner when the projection matrix is solved inexactly. In section 2, we first review the spectral properties of the preconditioned system when the preconditioners are constructed with the exact coarse space, then we estimate the spectral bounds of the preconditioned systems when the approximate coarse space is used. Section 3 presents the perturbation analysis on the spectrum of the preconditioned system in the case that the projection matrix have some perturbation. Numerical results are reported in Section 4.

2. Coarse space spanned by the approximate coarse space. In this section, we briefly review the spectrum of the system preconditioned by using the exact coarse space, then based on these properties we derive the bounds of the spectrum of the system preconditioned by using the approximate coarse space. Let (λ_i, v_i) be an eigenpair of A and v_i be normalized. (v_i, \dots, v_n) is orthogonal since A is SPD. The spectral decomposition of A can be written as

$$(2.1) \quad A = (V, V_\perp) \begin{pmatrix} \Lambda & 0 \\ 0 & \Lambda_\perp \end{pmatrix} \begin{pmatrix} V^T \\ V_\perp^T \end{pmatrix},$$

where $\Lambda = \text{diag}\{\lambda_1, \dots, \lambda_r\}$, $\Lambda_\perp = \text{diag}\{\lambda_{r+1}, \dots, \lambda_n\}$, $V = (v_1, \dots, v_r)$, and $V_\perp = (v_{r+1}, \dots, v_n)$.

Assume that $\lambda_1, \lambda_2, \dots, \lambda_r$ are small eigenvalues that impacts the convergence of the Krylov subspace methods. Theorem 2.1 and Theorem 2.2 show that if we take $Z = V$, the small eigenvalues are removed or shifted when the preconditioners are applied on either side of A , moreover, the rest of eigenvalues and all the eigenvectors are not changed.

THEOREM 2.1. Let $\tilde{E} = V^T A V$. Define

$$\begin{aligned}\tilde{P}_D &= I - A V \tilde{E}^{-1} V^T, \\ \tilde{P}_C &= I + V \tilde{E}^{-1} V^T, \\ \tilde{P}_A &= I - A V \tilde{E}^{-1} V^T + V \tilde{E}^{-1} V^T.\end{aligned}$$

Then we have the following spectral decomposition

$$\begin{aligned}\tilde{P}_D A &= (V, V_\perp) \begin{pmatrix} 0 & 0 \\ 0 & \Lambda_\perp \end{pmatrix} \begin{pmatrix} V^T \\ V_\perp^T \end{pmatrix}, \\ \tilde{P}_C A &= (V, V_\perp) \begin{pmatrix} I + \Lambda & 0 \\ 0 & \Lambda_\perp \end{pmatrix} \begin{pmatrix} V^T \\ V_\perp^T \end{pmatrix}, \\ \tilde{P}_A A &= (V, V_\perp) \begin{pmatrix} I & 0 \\ 0 & \Lambda_\perp \end{pmatrix} \begin{pmatrix} V^T \\ V_\perp^T \end{pmatrix}.\end{aligned}$$

Proof. From the definition of \tilde{E} , we have

$$\tilde{E} = V^T A V = V^T V \Lambda = \Lambda.$$

Next, let us consider the first result. Obviously, $\tilde{P}_D A V = 0$. We then have

$$\begin{aligned}\tilde{P}_D A V_\perp &= A V_\perp - A V \Lambda^{-1} V^T A V_\perp \\ &= V_\perp \Lambda_\perp - V V^T V_\perp \Lambda_\perp \\ &= V_\perp \Lambda_\perp.\end{aligned}$$

Thus the first spectral decomposition holds. Since

$$\tilde{P}_C A = A + A V \tilde{E}^{-1} V^T = A + V \Lambda \Lambda^{-1} V^T = A + V V^T,$$

we have

$$\tilde{P}_C A V = A V + V V^T V = V(\Lambda + I)$$

and

$$\tilde{P}_C A V_\perp = A V_\perp + V V^T V_\perp = A V_\perp = V_\perp \Lambda_\perp.$$

Hence, the second spectral decomposition is true. The third spectral decomposition follows from

$$\tilde{P}_A A V = A V - A V \tilde{E}^{-1} V^T A V + V \tilde{E}^{-1} V^T A V = V$$

and

$$\begin{aligned}\tilde{P}_A A V_\perp &= A V_\perp - A V \tilde{E}^{-1} V^T A V_\perp + V \tilde{E}^{-1} V^T A V_\perp \\ &= A V_\perp - A V \tilde{E}^{-1} V^T V_\perp \Lambda_\perp + V \tilde{E}^{-1} V^T V_\perp \Lambda_\perp \\ &= A V_\perp = V_\perp \Lambda_\perp\end{aligned}$$

It follows immediately from these three spectral decomposition that P_DA , P_CA and P_AA are symmetric. \square

THEOREM 2.2. \tilde{P}_D , \tilde{P}_C and \tilde{P}_A are defined in Theorem 2.1. Then we have $A\tilde{P}_D = \tilde{P}_DA$, $A\tilde{P}_C = \tilde{P}_CA$ and $A\tilde{P}_A = \tilde{P}_AA$.

Proof. The first result follows from

$$A\tilde{P}_DV = AV - AV\Lambda\Lambda^{-1}V^TV = 0 = \tilde{P}_DAV$$

and

$$A\tilde{P}_DV_\perp = AV_\perp + AV\Lambda^{-1}V^TV_\perp = V_\perp\Lambda_\perp = \tilde{P}_DAV_\perp.$$

Note that A , \tilde{P}_C and \tilde{P}_CA are symmetric. We then have

$$\tilde{P}_CA = (\tilde{P}_CA)^T = A^T\tilde{P}_C^T = A\tilde{P}_C.$$

Since

$$\begin{aligned} A\tilde{P}_A &= A - A^2V\tilde{E}^{-1}V^T + AV\tilde{E}^{-1}V^T \\ &= A - AV\Lambda\Lambda^{-1}V^T + V\Lambda\Lambda^{-1}V^T \\ &= A - AVV^T + VV^T, \end{aligned}$$

we get

$$A\tilde{P}_AV = AV - AVV^TV + VV^TV = V = \tilde{P}_AAV$$

and

$$A\tilde{P}_AV_\perp = AV_\perp + AVV^TV_\perp + VV^TV_\perp = AV_\perp = V_\perp\Lambda_\perp = \tilde{P}_AAV_\perp.$$

The third result is therefore true. In addition, AP_D , AP_C and AP_A are symmetric as well. \square

For a large system, it is impractical to build the preconditioners by using the exact eigenvectors associated with the small eigenvalues, since in general computing these eigenvectors is more costly and more difficult than solving a linear system. Nevertheless, for some cases, the approximate eigenvectors are cheaply obtained and thus are used to produce the preconditioners. For example, in the Newton method for solving nonlinear problems, the information obtained during solving the first linear system is reused to build the coarse spaces for accelerating the convergence of the succeeding linear systems, see [7] and [8]. We hope that the system preconditioned by using the approximate coarse space has the similar eigenvalue distribution as described in Theorem 2.1. This motivates us to analyse how the perturbation in the coarse space impacts the spectrum of the system preconditioned by P_D , P_C and P_A .

In the following discussion we assume that Z is column orthogonal. Let \mathcal{Z} be the subspace spanned by the columns of Z . Let \mathcal{Z}^\perp be the orthogonal complement of \mathcal{Z} and Z_\perp be an orthogonal basis of \mathcal{Z}^\perp . Likewise, let \mathcal{V} be the subspace spanned by the columns of V , \mathcal{V}^\perp be the orthogonal complement of \mathcal{V} and V_\perp be an orthogonal basis of \mathcal{V}^\perp . Let σ denote the singular value of a matrix. Let $\text{dist}(\mathcal{Z}, \mathcal{V})$ denote the distance between subspaces \mathcal{Z} and \mathcal{V} . It is shown in [6] that $\text{dist}(\mathcal{Z}, \mathcal{V})$ can be evaluated by either $\sigma_{\max}(Z^TV_\perp)$ or $\sigma_{\max}(V^TZ_\perp)$. Note that $0 \leq \text{dist}(\mathcal{Z}, \mathcal{V}) \leq 1$ since \mathcal{Z} and \mathcal{V} have the same dimension. We define the acute angle between subspaces \mathcal{Z} and \mathcal{V} as

$\theta = \arcsin \text{dist}(\mathcal{Z}, \mathcal{V})$. The next lemma shows that $\cos \theta$ can be evaluated in a similar way as $\sin \theta$.

LEMMA 2.3. *Let θ be the acute angle between subspaces \mathcal{Z} and \mathcal{V} that have the same dimension. Let Z and V be the orthogonal bases of \mathcal{Z} and \mathcal{V} respectively. Then*

$$\begin{aligned}\sin \theta &= \sigma_{\max}(Z^T V_{\perp}) = \sigma_{\max}(V^T Z_{\perp}), \\ \cos \theta &= \sigma_{\min}(Z^T V) = \sigma_{\min}(Z_{\perp}^T V_{\perp}).\end{aligned}$$

Proof. The first identity and its proof can be found in [6, Theorem 2.6.1]. We only prove the second identity here. Since (V, V_{\perp}) is orthogonal, it follows from $\|(V, V_{\perp})^T Zx\|_2 = 1$ for all unit 2-norm $x \in \mathbb{R}^r$ that $\|V^T Zx\|_2^2 + \|V_{\perp}^T Zx\|_2^2 = 1$. Thus

$$\begin{aligned}\sigma_{\min}(V^T Z)^2 &= \min_{\|x\|_2=1} \|V^T Zx\|_2^2 = 1 - \max_{\|x\|_2=1} \|V_{\perp}^T Zx\|_2^2 \\ &= 1 - \sigma_{\max}(V_{\perp}^T Z)^2 = \cos^2 \theta.\end{aligned}$$

Similarly, since (Z, Z_{\perp}) is orthogonal, it follows from $\|(Z, Z_{\perp})^T V_{\perp}x\|_2 = 1$ for all unit 2-norm $x \in \mathbb{R}^{(n-r)}$ that $\|Z^T V_{\perp}x\|_2^2 + \|Z_{\perp}^T V_{\perp}x\|_2^2 = 1$. Thus

$$\begin{aligned}\sigma_{\min}(Z_{\perp}^T V_{\perp})^2 &= \min_{\|x\|_2=1} \|Z_{\perp}^T V_{\perp}x\|_2^2 = 1 - \max_{\|x\|_2=1} \|Z^T V_{\perp}x\|_2^2 \\ &= 1 - \sigma_{\max}(Z^T V_{\perp})^2 = \cos^2 \theta.\end{aligned}$$

The second identity is therefore valid. \square

Let P_D be defined by (1.2). $P_D A$ is an symmetric matrix since A is SPD. Hence, Courant-Fischer Minimax Theorem [6] can be applied to estimate the eigenvalues of $P_D A$. As is known that if $A \in \mathbb{R}^{n \times n}$ is symmetric, then

$$\lambda_k(A) = \max_{\dim(S)=k} \min_{x \in S, \|x\|_2=1} x^T A x, \quad k = 1, \dots, n.$$

THEOREM 2.4. *Let P_D be defined by (1.2) and θ be the acute angle between \mathcal{V} and \mathcal{Z} . Then $P_D A$ has r zero eigenvalues. Moreover, if $\cos \theta \neq 0$, then the nonzero eigenvalues of $P_D A$ satisfy*

$$\lambda_{\min}(\Lambda_{\perp}) - \varepsilon_D \leq \lambda(P_D A) \leq \lambda_{\max}(\Lambda_{\perp}) + \eta_D,$$

where $\eta_D = \lambda_{\max}(\Lambda_{\perp})(\sin \theta + \sin^2 \theta)$ and $\varepsilon_D = \eta_D + \|E^{-1}\|_2(\|E\|_2 + \lambda_{\max}(\Lambda_{\perp}))^2 \tan^2 \theta$.

Proof. It follows from $P_D A Z = 0$ that $P_D A$ has r zero eigenvalues. From Theorem 2.1, we have

$$(2.2) \quad x^T \tilde{P}_D A x = x^T V_{\perp} \Lambda_{\perp} V_{\perp}^T x.$$

For all unit 2-norm $x \in \mathbb{R}^n$, it can be written as $x = x_1 + x_2$, where $x_1 \in \mathcal{Z}$ and $x_2 \in \mathcal{Z}^{\perp}$. Moreover, there exist $t \in \mathbb{R}^r$ and $s \in \mathbb{R}^{n-r}$ such that $x_1 = Zt$ and $x_2 = Z_{\perp}s$, and

$$\begin{aligned}x^T A Z E^{-1} Z^T A x &= x_1^T A Z E^{-1} Z^T A x_1 + x_1^T A Z E^{-1} Z^T A x_2 \\ &\quad + x_2^T A Z E^{-1} Z^T A x_1 + x_2^T A Z E^{-1} Z^T A x_2 \\ &= t^T Z^T A Z t + t^T Z^T A x_2 + x_2^T A Z t \\ &\quad + x_2^T A Z E^{-1} Z^T A x_2 \\ &= x_1^T A x_1 + x_1^T A x_2 + x_2^T A x_1 + x_2^T A Z E^{-1} Z^T A x_2 \\ &= x^T A x - x_2^T A x_2 + x_2^T A Z E^{-1} Z^T A x_2.\end{aligned}$$

Then we obtain

$$(2.3) \quad \begin{aligned} x^T P_D A x &= x^T A x - x^T A Z E^{-1} Z^T A x \\ &= x_2^T A x_2 - x_2^T A Z E^{-1} Z^T A x_2. \end{aligned}$$

Subtract (2.2) from (2.3) on both sides, we get

$$(2.4) \quad x^T P_D A x = x^T \tilde{P}_D A x + x_2^T A x_2 - x^T V_\perp \Lambda_\perp V_\perp^T x - x_2^T A Z E^{-1} Z^T A x_2.$$

The middle terms on the right-hand side of the above expression can be replaced by

$$\begin{aligned} x_2^T A x_2 - x^T V_\perp \Lambda_\perp V_\perp^T x &= x_2^T A x_2 - (x_1^T + x_2^T) V_\perp \Lambda_\perp V_\perp^T (x_1 + x_2) \\ &= x_2^T V \Lambda V^T x_2 - x_1^T V_\perp \Lambda_\perp V_\perp^T x_1 \\ &\quad - x_1^T V_\perp \Lambda_\perp V_\perp^T x_2 - x_2^T V_\perp \Lambda_\perp V_\perp^T x_1 \\ &= s^T (Z_\perp^T V) \Lambda (V^T Z_\perp) s - t^T (Z_\perp^T V_\perp) \Lambda_\perp (V_\perp^T Z) t \\ &\quad - t^T (Z_\perp^T V_\perp) \Lambda_\perp (V_\perp^T Z_\perp) s - s^T (Z_\perp^T V_\perp) \Lambda_\perp (V_\perp^T Z) t. \end{aligned}$$

$\|x_1\|_2 = \|t\|_2$ and $\|x_2\|_2 = \|s\|_2$ since Z and Z_\perp are column orthogonal. Using Lemma 2.3, we obtain

$$(2.5) \quad \begin{aligned} |x_2^T A x_2 - x^T V_\perp \Lambda_\perp V_\perp^T x| &\leq (\|x_2\|_2^2 \|\Lambda\|_2 + \|x_1\|_2^2 \|\Lambda_\perp\|_2) \sin^2 \theta \\ &\quad + 2\|x_1\|_2 \|x_2\|_2 \|\Lambda_\perp\|_2 \sin \theta \cos \theta. \end{aligned}$$

For the last item on the right-hand side of (2.4), we only need to estimate the bound of $Z_\perp^T A Z$ because of $x_2^T A Z E^{-1} Z^T A x_2 = s^T (Z_\perp^T A Z) E^{-1} (Z_\perp^T A Z)^T s$. It follows from $(Z, Z_\perp)(Z, Z_\perp)^T = I$ that

$$A Z = Z (Z^T A Z) + Z_\perp (Z_\perp^T A Z).$$

Multiply by V_\perp^T from the left on both sides of the above equation, we have

$$\begin{aligned} (V_\perp^T Z_\perp)(Z_\perp^T A Z) &= V_\perp^T A Z - (V_\perp^T Z)(Z^T A Z) \\ &= \Lambda_\perp (V_\perp^T Z) - (V_\perp^T Z)(Z^T A Z). \end{aligned}$$

$V_\perp^T Z_\perp$ is invertible since $\cos \theta \neq 0$. Thus

$$(2.6) \quad (Z_\perp^T A Z) = (V_\perp^T Z_\perp)^{-1} \Lambda_\perp (V_\perp^T Z) - (V_\perp^T Z_\perp)^{-1} (V_\perp^T Z) E.$$

Using Lemma 2.3, we have $\|Z_\perp^T A Z\|_2 \leq (\|E\|_2 + \|\Lambda_\perp\|_2) \tan \theta$. Hence

$$(2.7) \quad x_2^T A Z E^{-1} Z^T A x_2 \leq \|x_2\|_2^2 \|E^{-1}\|_2 (\|E\|_2 + \|\Lambda_\perp\|_2)^2 \tan^2 \theta.$$

E^{-1} is SPD since A is SPD. Consequently,

$$(2.8) \quad x_2^T A Z E^{-1} Z^T A x_2 = (Z^T A x_2)^T E^{-1} (Z^T A x_2) \geq 0.$$

$P_D A$ is symmetric since A is SPD. Applying Courant-Fischer Minimax Theorem to (2.4) with (2.5) and (2.8), we have

$$\begin{aligned} \lambda(P_D A) &\leq \lambda(\tilde{P}_D A) + |x_2^T A x_2 - x^T V_\perp \Lambda_\perp V_\perp^T x| \\ &\leq \lambda_{\max}(\Lambda_\perp) + \lambda_{\max}(\Lambda_\perp)(\sin \theta + \sin^2 \theta). \end{aligned}$$

Note that P_DA has r zero eigenvalues. Again, applying Courant-Fischer Minimax Theorem to (2.4) with (2.5) and (2.7), the lower bound of the nonzero eigenvalues of P_DA is given by

$$\begin{aligned}\lambda(P_DA) &\geq \lambda(\tilde{P}_DA) - |x_2^T Ax_2 - x^T V_\perp \Lambda_\perp V_\perp^T x| - x_2^T AZE^{-1}Z^T Ax_2 \\ &\geq \lambda_{\min}(\Lambda_\perp) - \lambda_{\max}(\Lambda_\perp)(\sin \theta + \sin^2 \theta) \\ &\quad - \|E^{-1}\|_2(\|E\|_2 + \lambda_{\max}(\Lambda_\perp))^2 \tan^2 \theta.\end{aligned}$$

As a result the theorem is true. \square

The above theorem shows that in exact arithmetic, as θ approaches zero, the maximal and minimal nonzero eigenvalues of P_DA converge to $\lambda_{\max}(\Lambda_\perp)$ and $\lambda_{\min}(\Lambda_\perp)$, respectively. Hence with an appropriate coarse space the spectrum of P_DA is similar to that of \tilde{P}_DA . When there exists rounding error, however, P_DAZ may not be equal to a zero matrix. In this case, P_DA possibly has some eigenvalues around zero that should be equal to zero in exact arithmetic. So there is a potential risk for P_D to yield a poor spectrum of the preconditioned system.

The authors in [19] investigated the properties of P_D , P_{BNN} and P_{ADEF1} . They established the relations between these preconditioners in terms of the spectrum. Suppose that M is an SPD matrix. Let the spectrum of $P_DM^{-1}A$ be given by $\{0, \dots, 0, \gamma_{r+1}, \dots, \gamma_n\}$ with $\gamma_{r+1} \leq \gamma_{r+2} \leq \dots \leq \gamma_n$. Let the spectrum of $P_{BNN}A$ and $P_{ADEF1}A$ be $\{1, \dots, 1, \mu_{r+1}, \dots, \mu_n\}$ with $\mu_{r+1} \leq \mu_{r+2} \leq \dots \leq \mu_n$. Then, $\gamma_i = \mu_i$ for all $i = r+1, \dots, n$. The proof of this result can be found in [19, Theorem 3.3]. From the relation of P_D and P_{ADEF1} , we immediately obtain the next corollary if we take $M = I$.

COROLLARY 2.5. *Let the spectrum of P_DA be given by $\{0, \dots, 0, \tilde{\lambda}_{r+1}, \dots, \tilde{\lambda}_n\}$. Then the spectrum of P_AA is $\{1, \dots, 1, \tilde{\lambda}_{r+1}, \dots, \tilde{\lambda}_n\}$.*

Corollary 2.5 implies that if the spectrum of P_DA is known, the one of P_AA would be known. The spectral bounds of P_DA are described in Theorem 2.4, so we can easily bound the spectrum of P_AA .

THEOREM 2.6. *Let P_A be defined by (1.5). Let θ the acute angle between subspaces \mathcal{Z} and \mathcal{V} . If $\cos \theta \neq 0$, then the eigenvalues of P_AA satisfy*

$$\min\{1, \lambda_{\min}(\Lambda_\perp) - \varepsilon_D\} \leq \lambda(P_AA) \leq \max\{1, \lambda_{\max}(\Lambda_\perp) + \eta_D\},$$

where η_D and ε_D are defined in Theorem 2.4.

Proof. It follows directly from Theorem 2.4 and Corollary 2.5. \square

Next, we consider the spectrum of P_CA . Note that P_CA is not necessarily symmetric although both A and P_C are symmetric. So we can not apply Courant-Fischer Minimax theorem to estimate the eigenvalues of P_CA , which are positive since P_CA is similar to a SPD matrix, and are a subset of $x^T P_CA x$ for all unit 2-norm $x \in \mathbb{R}^n$. Hence the eigenvalues of P_CA can be bounded by estimating $x^T P_CA x$.

THEOREM 2.7. *Let P_C be defined by (1.3), and θ the acute angle between \mathcal{V} and \mathcal{Z} . If $\cos \theta \neq 0$, then*

$$\begin{aligned}\lambda_{\max}(P_CA) &\leq \max\{1 + \lambda_{\max}(\Lambda), \lambda_{\max}(\Lambda_\perp)\} + \varepsilon_C, \\ \lambda_{\min}(P_CA) &\geq \min\{1 + \lambda_{\min}(\Lambda), \lambda_{\min}(\Lambda_\perp)\} - \varepsilon_C,\end{aligned}$$

where $\varepsilon_C = \frac{1}{2}(\lambda_{\max}(\Lambda_\perp)\|E^{-1}\|_2 + 1)\tan \theta + \sin \theta + \sin^2 \theta$.

Proof. P_CA is similar to $A + A^{1/2}ZE^{-1}Z^T A^{1/2}$ since A is SPD. Moreover $A + A^{1/2}ZE^{-1}Z^T A^{1/2}$ is SPD as well since A and E are SPD. Thus the eigenvalues of P_CA are positive.

For all unit 2-norm $x \in \mathbb{R}^n$, we write $x = x_1 + x_2$, where $x_1 \in \mathcal{Z}$ and $x_2 \in \mathcal{Z}_\perp$. There exists $t \in \mathbb{R}^r$ such that $x_1 = Zt$, likewise, there is $s \in \mathbb{R}^{n-r}$ such that $x_2 = Z_\perp s$.

Then $x^T P_C A x$ can be expressed as follows

$$\begin{aligned}
 (2.9) \quad x^T P_C A x &= x^T A x + (x_1 + x_2)^T Z E^{-1} Z^T A (x_1 + x_2) \\
 &= x^T A x + x_1^T Z E^{-1} Z^T A x_1 + x_1^T Z E^{-1} Z^T A x_2 \\
 &= x^T A x + x_1^T x_1 + x_1^T Z E^{-1} Z^T A x_2.
 \end{aligned}$$

From the definition of \tilde{P}_C in (2.1), we obtain

$$\begin{aligned}
 (2.10) \quad x^T \tilde{P}_C A x &= x^T A x + x^T V \tilde{E}^{-1} V^T A x \\
 &= x^T A x + x^T V V^T x.
 \end{aligned}$$

Subtract (2.10) from (2.9), we have

$$\begin{aligned}
 (2.11) \quad x^T P_C A x - x^T \tilde{P}_C A x &= x_1^T Z E^{-1} Z^T A x_2 + x_1^T x_1 - x^T V V^T x \\
 &= x_1^T Z E^{-1} Z^T A x_2 + x_1^T V_\perp V_\perp^T x_1 \\
 &\quad - x_2^T V V^T x_2 - 2x_1^T V V^T x_2.
 \end{aligned}$$

Since both A and E^{-1} are symmetric,

$$x_1^T Z E^{-1} Z^T A x_2 = x_2^T A Z E^{-1} Z^T x_1 = s_2^T Z_\perp^T A Z E^{-1} t.$$

Using (2.6), we have

$$(Z_\perp^T A Z) E^{-1} = (V_\perp^T Z_\perp)^{-1} \Lambda_\perp (V_\perp^T Z) E^{-1} - (V_\perp^T Z_\perp)^{-1} (V_\perp^T Z).$$

Note that $\|x_1\|_2 = \|t\|_2$ and $\|x_2\|_2 = \|s\|_2$. Using Lemma 2.3, we obtain

$$(2.12) \quad |x_1^T Z E^{-1} Z^T A x_2| \leq \|x_1\|_2 \|x_2\|_2 (\|\Lambda_\perp\|_2 \|E^{-1}\|_2 + 1) \tan \theta.$$

Since $\|x\|_2 = 1$, we have the following bounds with Lemma 2.3

$$\begin{aligned}
 (2.13) \quad 0 &\leq x_1^T V_\perp V_\perp^T x_1 \leq \|x_1\|_2^2 \sin^2 \theta, \\
 0 &\leq x_2^T V V^T x_2 \leq \|x_2\|_2^2 \sin^2 \theta, \\
 |x_1^T V V^T x_2| &\leq \|x_1\|_2 \|x_2\|_2 \sin \theta.
 \end{aligned}$$

With (2.11), (2.12) and (2.13), we obtain

$$\lambda_{\min}(\tilde{P}_C A) - \varepsilon_C \leq x^T P_C A x \leq \lambda_{\max}(\tilde{P}_C A) + \varepsilon_C,$$

where $\varepsilon_C = \frac{1}{2}(\lambda_{\max}(\Lambda_\perp) \|E^{-1}\|_2 + 1) \tan \theta + \sin \theta + \sin^2 \theta$. Thus the theorem follows from $\min\{x^T P_C A x\} \leq \lambda(P_C A) \leq \max\{x^T P_C A x\}$ for all unit 2-norm $x \in \mathbb{R}^n$. \square

From Theorem 2.7, we conclude that the maximal and minimal eigenvalues of $P_C A$ converge to $\max\{\lambda_{\max}(\Lambda_\perp), 1 + \lambda_{\max}(\Lambda)\}$ and $\min\{\lambda_{\min}(\Lambda_\perp), 1 + \lambda_{\min}(\Lambda)\}$ respectively as θ approaches zero. With an appropriate coarse space, the spectral distribution of $P_C A$ would be close to that of $P_C A$. Analogously, it can be proved that the maximal and minimal eigenvalues of $A P_C$ have the same bounds as that of $P_C A$.

3. Inexact inverse of projection matrix. In practice, we do not form E^{-1} explicitly to compute $y = E^{-1}x$. Here y and x are vectors with the suitable size. Instead, we compute the LU factorization of E once, then solve the two triangular linear systems to obtain y . But it is expensive to compute the LU factorization when the matrix E is large. We therefore replace E by a perturbed one that is cheaper to compute. In this section, we analyse how the perturbation in the projection matrix impacts the spectrum of the preconditioned matrix. Assume \tilde{H} is an invertible matrix as an approximation to \tilde{E} in (2.1).

THEOREM 3.1. *Let $\bar{P}_D = I - AV\tilde{H}^{-1}V^T$ and $\rho = \tilde{E}\tilde{H}^{-1} - I$. If the eigenvalues of $\bar{P}_D A$ are real, then*

$$-\xi_D \leq \lambda(\bar{P}_D A) \leq \lambda_{\max}(\Lambda_{\perp}) + \xi_D,$$

where $\xi_D = \|\rho\|_2 \|\Lambda\|_2$.

Proof. $\|\rho\|_2$ measures how much \tilde{H} is approximate to \tilde{E} when \tilde{H}^{-1} is used as the right inverse of \tilde{E} . Obviously $\|\rho\|_2 = 0$ if and only if $\tilde{E} = \tilde{H}$. Assume $\|x\|_2 = 1$ and use the definition of \bar{P}_D in Theorem 2.1, we obtain

$$\begin{aligned} x^T \bar{P}_D A x &= x^T A x - x^T A V \tilde{H}^{-1} V^T A x \\ &= x^T A x - x^T V V^T A x - x^T V \rho V^T A x \\ &= x^T \tilde{P}_D A x - x^T V \rho \Lambda V^T x. \end{aligned}$$

Since $|x^T V \rho \Lambda V^T x| \leq \|\rho\|_2 \|\Lambda\|_2$,

$$-\|\rho\|_2 \|\Lambda\|_2 \leq x^T \bar{P}_D A x \leq \lambda_{\max}(\tilde{P}_D A) + \|\rho\|_2 \|\Lambda\|_2.$$

Since the eigenvalues of $\bar{P}_D A$ are real, the theorem follows from $\min\{x^T \bar{P}_D A x\} \leq \lambda(\bar{P}_D A) \leq \max\{x^T \bar{P}_D A x\}$. \square

Theorem 3.1 states that $\bar{P}_D A$ might have small eigenvalues around zero if $\|\rho\|_2 \neq 0$, which leads to the worse spectral distribution than that of $\tilde{P}_D A$.

THEOREM 3.2. *Let $\bar{P}_C = I + V\tilde{H}^{-1}V^T$ and $\rho = \tilde{H}^{-1}\tilde{E} - I$. If the eigenvalues of $\bar{P}_C A$ are real, then*

$$\begin{aligned} \lambda_{\max}(\bar{P}_C A) &\leq \max\{1 + \lambda_{\max}(\Lambda), \lambda_{\max}(\Lambda_{\perp})\} + \xi_C, \\ \lambda_{\min}(\bar{P}_C A) &\geq \min\{1 + \lambda_{\min}(\Lambda), \lambda_{\min}(\Lambda_{\perp})\} - \xi_C, \end{aligned}$$

where $\xi_C = \|\rho\|_2$.

Proof. In this theorem, we use \tilde{H}^{-1} as the left inverse of \tilde{E} . Assume $\|x\|_2 = 1$ and use the definition of \bar{P}_C in Theorem 2.1, then we have

$$\begin{aligned} x^T \bar{P}_C A x &= x^T A x + x^T V \tilde{H}^{-1} V^T A x \\ &= x^T A x + x^T V \tilde{E}^{-1} V^T A x + x^T V \rho V^T x \\ &= x^T \tilde{P}_C A x + x^T V \rho V^T x. \end{aligned}$$

Since $|x^T V \rho V^T x| \leq \|\rho\|_2$,

$$\lambda_{\min}(\tilde{P}_C A) - \|\rho\|_2 \leq x^T \bar{P}_C A x \leq \lambda_{\max}(\tilde{P}_C A) + \|\rho\|_2.$$

Since the eigenvalues of $\bar{P}_C A$ are real, the theorem follows from $\min\{x^T \bar{P}_C A x\} \leq \lambda(\bar{P}_C A) \leq \max\{x^T \bar{P}_C A x\}$. \square

Theorem 3.2 implies that if the eigenvalues of $\bar{P}_C A$ are real, the spectral distribution of $\bar{P}_C A$ is a little influenced with small $\|\rho\|_2$, because the maximal and minimal

eigenvalues of $\bar{P}_C A$ converge to $\max\{\lambda_{\max}(\Lambda_\perp), 1 + \lambda_{\max}(\Lambda)\}$ and $\min\{\lambda_{\min}(\Lambda_\perp), 1 + \lambda_{\min}(\Lambda)\}$ respectively as ρ approaches zero.

In next theorem, we need to estimate the difference of \tilde{H} and \tilde{E} when \tilde{H}^{-1} is used as both the left and the right inverse of \tilde{E} .

THEOREM 3.3. *Let $\bar{P}_A = I - AV\tilde{H}^{-1}V^T + V\tilde{H}^{-1}V^T$. Let $\rho_1 = \tilde{E}\tilde{H}^{-1} - I$ and $\rho_2 = \tilde{H}^{-1}\tilde{E} - I$. If the eigenvalues of $\bar{P}_A A$ are real, then*

$$\begin{aligned}\lambda_{\max}(\bar{P}_A A) &\leq \max\{1, \lambda_{\max}(\Lambda_\perp)\} + \xi_A, \\ \lambda_{\min}(\bar{P}_A A) &\geq \min\{1, \lambda_{\min}(\Lambda_\perp)\} - \xi_A,\end{aligned}$$

where $\xi_A = \|\rho_1\|_2 \|\Lambda\|_2 + \|\rho_2\|_2$.

Proof. Assume $\|x\|_2 = 1$ and use the definition of \tilde{P}_A in (2.1), we have

$$\begin{aligned}x^T \bar{P}_A A x &= x^T A x - x^T A V \tilde{H}^{-1} V^T A x + x^T V \tilde{H}^{-1} V^T A x \\ &= x^T A x - x^T V \rho_1 \Lambda V^T x + x^T V \rho_2 V^T x \\ &\quad - x^T A V \tilde{E}^{-1} V^T A x + x^T V \tilde{E}^{-1} V^T A x \\ &= x^T \tilde{P}_A A x - x^T V \rho_1 \Lambda V^T x + x^T V \rho_2 V^T x.\end{aligned}$$

Since $|x^T V \rho_1 \Lambda V^T x - x^T V \rho_2 V^T x| \leq \|\rho_1\|_2 \|\Lambda\|_2 + \|\rho_2\|_2$,

$$\lambda_{\min}(\tilde{P}_A A) - \|\rho_1\|_2 \|\Lambda\|_2 - \|\rho_2\|_2 \leq x^T \bar{P}_A A x \leq \lambda_{\max}(\tilde{P}_A A) + \|\rho_1\|_2 \|\Lambda\|_2 + \|\rho_2\|_2.$$

We assumed that the eigenvalues of $\bar{P}_A A$ are real. Thus the theorem follows from $\min\{x^T \bar{P}_A A x\} \leq \lambda(\bar{P}_A A) \leq \max\{x^T \bar{P}_A A x\}$. \square

4. Numerical experiments. In this section, a numerical comparison of various preconditioners is reported. All tests are performed with Matlab (R2010b) on an Intel Core2 Duo E7500, 2.93GHz processor with 4Gb memory. Except for the deflation preconditioner, the system preconditioned by the coarse correction and adapted deflation preconditioners are not necessarily symmetric although A is SPD. Therefore we apply GMRES [16] to solve the preconditioned system iteratively. In addition, we apply Gram-Schmidt method with reorthogonalization to maintain the orthogonality of basis of the Krylov subspace [6, 16].

4.1. Diagonal matrix. The first test case is a diagonal matrix with entries 10^{-7} , 10^{-6} , \dots , 10^{-1} , 1 , 10 , 10.1 , 10.2 , \dots , 209 , 209.1 . The matrix has 7 small eigenvalues less than 1 to be removed. The eigenvectors associated with these eigenvalues are the unit vectors, i.e., $V = (e_1, e_2, \dots, e_7)$, where e_i is the i th column of the identity matrix. The right-hand side is a vector of all ones. The initial guess vector is a zero vector. All tests are required to reduce the relative residual below 10^{-12} . GMRES method without preconditioning converges at 273th iteration. The perturbations in the coarse space and the projection matrix are generated by the Matlab function `rand`.

Table 4.1 shows the distance between the exact coarse space and the coarse space with various perturbations. It should be noted that it is expensive and unnecessary in practice to compute $\sin \theta$ by Lemma 2.3 for a general linear system. Assume that $(\tilde{\lambda}_i, \tilde{v}_i)$ ($i = 1, \dots, r$) are Ritz pairs that are extracted from the perturbed coarse space by the Rayleigh-Ritz procedure [18]. In general, $\max_{i=1, \dots, r} \{\|A\tilde{v}_i - \tilde{\lambda}_i \tilde{v}_i\|_2\}$ denoted by res_{\max} in Table 4.1 decreases as the two subspaces approach each other. So we can use res_{\max} to measure the distance between the two subspaces since it is more convenient to compute.

In Table 4.2, the second column shows the number of GMRES iterations with the three preconditioners in the case that there is no perturbation in the coarse space

TABLE 4.1
The distance between $\text{span}\{V\}$ and $\text{span}\{V + \text{rand}/\varepsilon\}$.

	$\varepsilon = 1e + 01$	$\varepsilon = 1e + 02$	$\varepsilon = 1e + 03$	$\varepsilon = 1e + 04$	$\varepsilon = 1e + 05$
$\sin \theta$	9.77e-01	5.06e-01	6.03e-02	6.05e-03	6.02e-04
res_{max}	7.08e+01	5.54e+01	7.33	5.78e-01	4.43e-02

and the projection matrix. The columns 3-7 show the number of GMRES iterations when only the coarse space has some perturbation. As is shown, all preconditioners suffer from the perturbation if it is large (see the third column). As the perturbation decreases, P_C and P_A become better, whereas P_D becomes better only when the perturbation is very small. On the other hand, if the perturbed coarse space is close enough to the exact one (see the last two columns), P_D is slightly more efficient than P_A , and both of them are more efficient than P_C .

TABLE 4.2
The number of GMRES iterations with various preconditioners that are constructed with $Z = V + \text{rand}/\varepsilon$ and E^{-1} .

	V, \tilde{E}^{-1}	$\varepsilon = 1e + 01$	$\varepsilon = 1e + 02$	$\varepsilon = 1e + 03$	$\varepsilon = 1e + 04$	$\varepsilon = 1e + 05$
P_D	71	>300	>300	>300	109	88
P_C	104	273	267	222	173	144
P_A	72	290	231	167	117	96

Figure 4.1 and Figure 4.2 show the eigenvalue distribution of the system preconditioned by the three preconditioners. As discussed in Section 2, because of rounding error, $P_D A$ has some tiny eigenvalues around zero. In general, it is difficult to figure out the condition under which these tiny eigenvalues cause the stagnation in the convergence. For the test case of diagonal matrix, rounding error does not impact the convergence of $P_D A$ when the perturbation in the coarse space is significantly small. We also see that the spectrum of $P_A A$ is more clustered than that of $P_C A$, which is consistent with the estimated bounds of their spectrum described in Theorem 2.6 and Theorem 2.7.

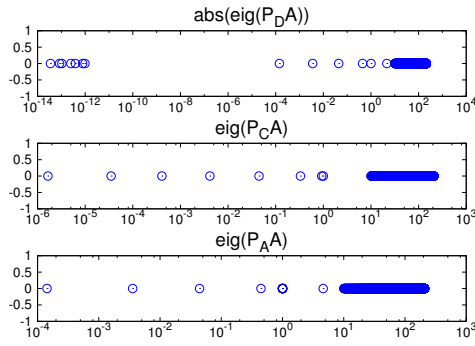


FIG. 4.1. The eigenvalue distribution of the preconditioned system. The preconditioners P_D , P_C and P_A are built with $Z = V + \text{rand}/1e + 03$ and E^{-1} .

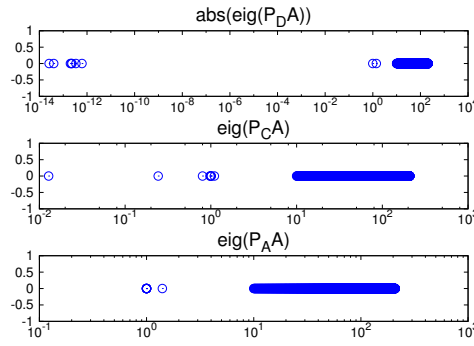


FIG. 4.2. The eigenvalue distribution of the preconditioned system. The preconditioners P_D , P_C and P_A are built with $Z = V + \text{rand}/1e + 05$ and E^{-1} .

In Table 4.3, the first three rows show the number of GMRES iterations when only the projection matrix has perturbation. The difference of \tilde{H}^{-1} and \tilde{E}^{-1} is reported

in the last two rows. In comparison with the second column in Table 4.2, we conclude the perturbation in the projection matrix has a little impact on P_C and P_A , but has a severe impact on P_D even when \tilde{H}^{-1} is very close to \tilde{E}^{-1} (see the last column).

TABLE 4.3

The number of GMRES iterations with different preconditioners, where only the projection matrix is perturbed and the perturbation of \tilde{E} is $\tilde{H} = \tilde{E} + \text{rand}/\varepsilon$.

	$\varepsilon = 1e + 10$	$\varepsilon = 1e + 12$	$\varepsilon = 1e + 14$	$\varepsilon = 1e + 16$
P_D	>300	>300	>300	>300
P_C	111	104	104	104
P_A	88	88	80	80
$\ \tilde{H}^{-1}\tilde{E} - I\ _2$	1.68e-03	1.08e-05	1.77e-07	1.23e-09
$\ \tilde{E}\tilde{H}^{-1} - I\ _2$	1.68e-03	1.08e-05	1.77e-07	1.23e-09

Figure 4.3 and Figure 4.4 report the spectral distribution of P_D , P_C and P_A . We see that P_C and P_A successfully shift the small eigenvalues of A to around one. Due to the perturbation in the projection matrix, P_D fails to deflate the small eigenvalues and thus yields some tiny eigenvalues around zero, which leads to a worse convergence (see the first row in Table 4.3).

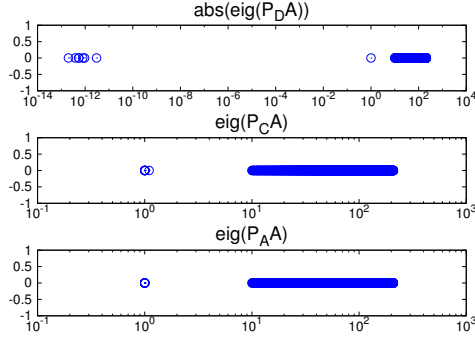


FIG. 4.3. The eigenvalue distribution of the preconditioned system, where the preconditioners P_D , P_C and P_A are built with V and $\tilde{H} = \tilde{E} + \text{rand}/1e + 12$.

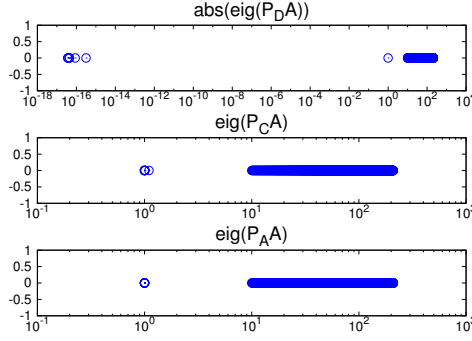


FIG. 4.4. The eigenvalue distribution of the preconditioned system, where the preconditioners P_D , P_C and P_A are built with V and $\tilde{H} = \tilde{E} + \text{rand}/1e + 16$.

4.2. Boundary value problem. We solve the following model problem

$$\begin{aligned} -\nabla \cdot (\kappa \nabla u) &= f \quad \text{in } \Omega = [0, 1]^2, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned}$$

by the two-level multiplicative Schwarz method [8]. Here, κ is the diffusion function of x and y . The model problem is discretized by FreeFem++ [15] and the resulting coefficient matrix is of size 10201. Tests are performed on irregular overlapping decompositions with the overlap of 2 elements. These overlapping decompositions are built by adding the immediate neighboring vertices to non-overlapping subdomain obtained by Metis [9].

In the two-level multiplicative Schwarz method, the first level preconditioner is the restricted additive Schwarz preconditioner (RAS) [1] that is responsible to remove high frequency modes of the original system, and the deflation, coarse correction and

adapted deflation preconditioners are applied as the second level preconditioners that remove lower frequency ones of the system preconditioned by one-level preconditioner [8, 11].

We choose Ritz vectors to span the coarse space, which are extracted from the Krylov subspace during the solve of the system preconditioned by RAS. These vectors are the approximate eigenvectors corresponding to the lower part of the spectrum of the preconditioned system. To enrich the information on lower part of the spectrum, we construct the coarse space by splitting Ritz-vectors, see [8, 11, 19] and references therein. More precisely, let

$$V = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1,r} \\ v_{21} & v_{22} & \cdots & v_{2,r} \\ \cdots & & & \\ v_{nparts,1} & v_{nparts,2} & \cdots & v_{nparts,r} \end{bmatrix}$$

store Ritz vectors columnwise, where $nparts$ is the number of subdomains and r the number of Ritz vectors; let Z_i store the orthogonal vectors obtained by orthogonalizing $(v_{i1}, v_{i2}, \dots, v_{i,r})$. Then Z is formed as follows

$$Z = \begin{bmatrix} Z_1 & 0 & \cdots & 0 \\ 0 & Z_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & Z_{nparts} \end{bmatrix}.$$

Obviously, $\text{span}\{V\}$ is a subspace of $\text{span}\{Z\}$; $\text{span}\{Z\}$ is $nparts$ times as large as $\text{span}\{V\}$. So $\text{span}\{Z\}$ might have richer information corresponding to small eigenvalues.

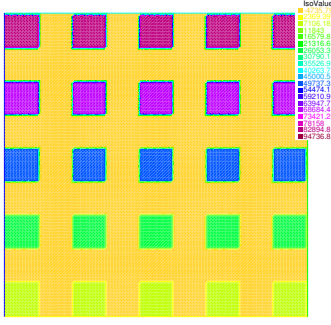
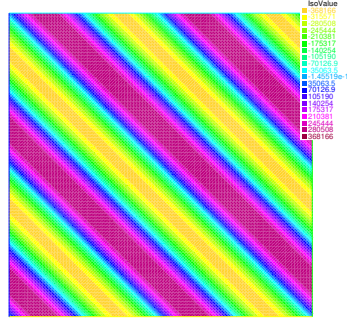
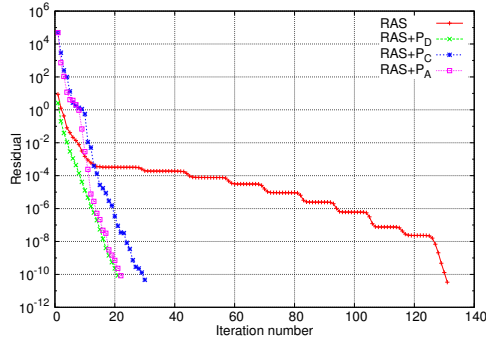
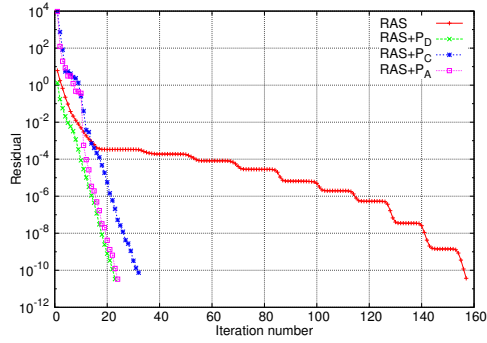
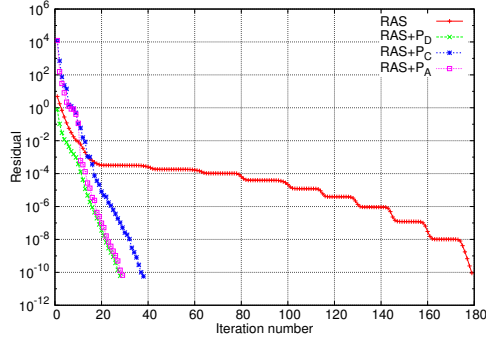
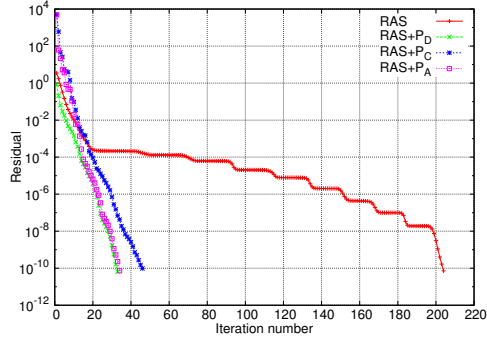
Some comparisons of the three preconditioners are performed on two different configurations with highly heterogeneous coefficient κ . See [11] for details. Two cases are described as following:

- skyscraper κ : for x and y such that for $[9x] \equiv 0 \pmod{2}$ and $[9y] \equiv 0 \pmod{2}$, $\kappa = 10^4([9y] + 1)$; and $\kappa = 1$ elsewhere. See Figure 4.5.
- continuous κ : $\kappa(x, y) = 10^6/3 \sin(4\pi(x + y) + 0.1)$. See Figure 4.6.

For both cases, we use the zero vector as the initial guess vector. The iteration will stop when the relative residual is less than 10^{-10} . Moreover, we construct the coarse space with all the eigenvectors associated with the eigenvalues less than 0.5 against the various domain decompositions.

In Figures 4.7-4.10, we report the convergence curves for the skyscraper case against the various number of subdomains. Compared to one-level method, all three two-level methods improve convergence sufficiently. P_D and P_A have almost the same number of iterations although the initial residual of P_D is much less than that of P_A . P_D and P_A are more efficient than P_C . Two-level method varies slightly on the number of iterations as the number of subdomains increases, while one-level method does.

Figures 4.11-4.14 plot the convergence curves for the continuous case against the various number of subdomains. Once again, we see that two-level method with RAS and the three preconditioners all outperform one-level method with only RAS. Like the skyscraper case, P_D and P_A have almost the same number of iterations for all four decompositions and outperform P_C .

FIG. 4.5. *Skyscraper case*FIG. 4.6. *Continuous case*FIG. 4.7. *Skyscraper case with 16 subdomains. 16 Ritz vectors spanning the coarse space.*FIG. 4.8. *Skyscraper case with 32 subdomains. 16 Ritz vectors spanning the coarse space.*FIG. 4.9. *Skyscraper case with 64 subdomains. 16 Ritz vectors spanning the coarse space.*FIG. 4.10. *Skyscraper case with 128 subdomains. 16 Ritz vectors spanning the coarse space.*

In Table 4.4, we report the maximal residual of Ritz pairs for both cases that are extracted from the Krylov subspace when solving the system preconditioned only by RAS.

Note that the projection matrix E is large in the case that the decomposition has 64 or 128 subdomains. As a consequence, computing LU factorization of E^{-1} is costly and impairs the gains in the number of iterations. Hence, we attempt to compute the incomplete LU factorization of E , which is cheaper to compute than the

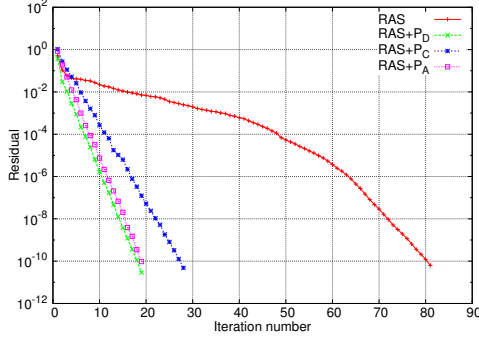


FIG. 4.11. Continuous case with 16 subdomains. 15 Ritz vectors spanning the coarse space.

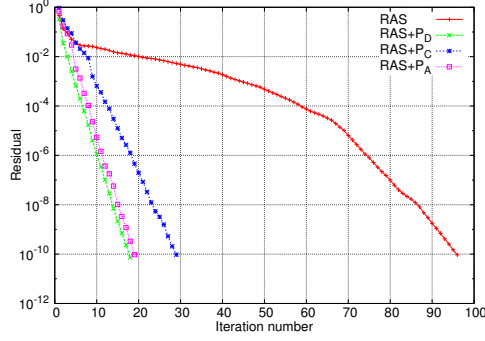


FIG. 4.12. Continuous case with 32 subdomains. 15 Ritz vectors spanning the coarse space.

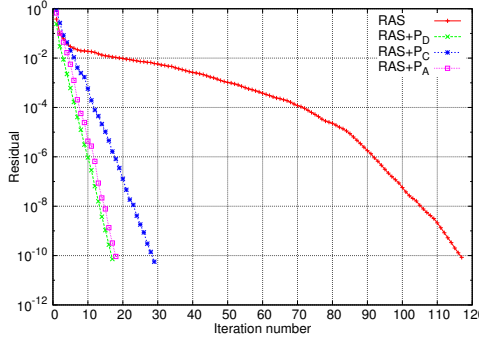


FIG. 4.13. Continuous case with 64 subdomains. 16 Ritz vectors spanning the coarse space.

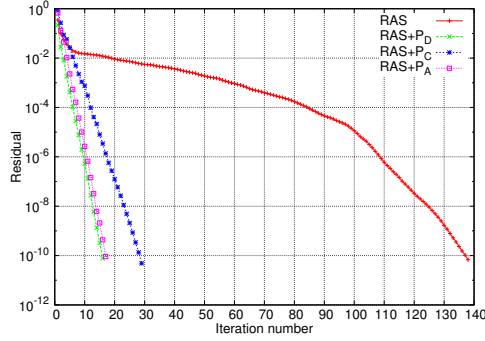


FIG. 4.14. Continuous case with 128 subdomains. 15 Ritz vectors spanning the coarse space.

TABLE 4.4

Maximal residual of Ritz pairs for the skyscraper and continuous cases.

Nparts	16	32	64	128
skyscraper	2.564719e-09	9.162259e-08	1.577898e-08	4.747077e-09
continuous	4.439121e-03	5.242112e-03	4.749042e-03	1.326165e-03

LU factorization. Assume L and U are factors of an incomplete LU factorization with no fill-in ($ILU(0)$) of E . Note that E has a sparse structure because of the sparse structure of Z . So L and U are sparse as well. This means that it is very cheap to solve $(LU)x = y$. In this way, E is actually replaced by LU .

From Figure 4.15 and Figure 4.16, it appears that the perturbation in E has almost no impact on P_A and P_C for the skyscraper case, but has severe impact on P_D that leads to a stagnation in the convergence. Figure 4.17 and Figure 4.18 show that P_A and P_C are also stable for the continuous case. As is shown in Figure 4.18, P_D is still unstable when the perturbation in E is not small enough (see the second column in Table 4.6). However, Figure 4.17 shows that P_D is stable, since the perturbation is small enough such that LU is almost same as E (see the first column in Table 4.6).

Table 4.5 and Table 4.6 present the distance between LU and E for both cases, respectively.

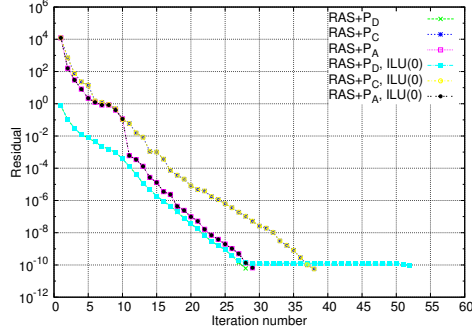


FIG. 4.15. *Skyscraper case with 64 subdomains. E is replaced by LU .*

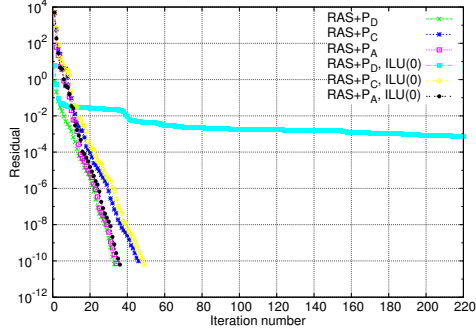


FIG. 4.16. *Skyscraper case with 128 subdomains. E is replaced by LU .*

TABLE 4.5
The distance between LU and E for the skyscraper case.

Nparts	64	128
$\ E(LU)^{-1} - I\ _2$	3.8588e-08	8.9712e+02
$\ (LU)^{-1}E - I\ _2$	4.1433e-10	8.6292

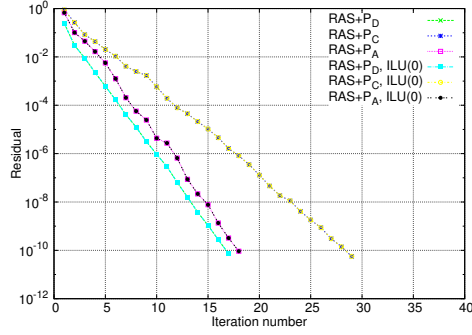


FIG. 4.17. *Continuous case with 64 subdomains. E is replaced by LU .*

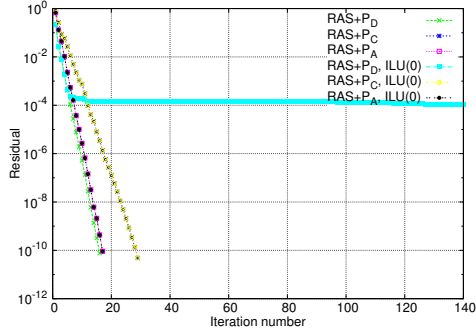


FIG. 4.18. *Continuous case with 128 subdomains. E is replaced by LU .*

TABLE 4.6
The distance between LU and E for the continuous case.

Nparts	64	128
$\ E(LU)^{-1} - I\ _2$	4.1008e-14	5.4090e-01
$\ (LU)^{-1}E - I\ _2$	1.3890e-15	1.0215e-01

5. Conclusion. We presented a perturbation analysis on the deflation, coarse correction and adapted deflation preconditioners when the inexact coarse space and inverse of projection matrix are applied for the construction of the preconditioners. Our analysis shows that in exact arithmetic the spectrum of the system preconditioned by the three preconditioners is impacted by the angle between the exact coarse space and the perturbed one. Moreover, we prove that the coarse correction and adapted deflation preconditioners are insensitive to the perturbation of the projection matrix, whereas the deflation preconditioner is sensitive. Numerical results of the different

test cases confirm the perturbation analysis.

Acknowledgements. The author would like to appreciate Professor Xiao-Chuan Cai for his valuable comments that are very helpful to improve the presentation of this paper.

References.

- [1] X.-C. CAI AND M. SARKIS, *A restricted additive Schwarz preconditioner for general sparse linear systems*, SIAM J. Sci. Comput., 21 (1999), pp. 792-797.
- [2] Y. A. ERLANGGA AND R. NABBEN, *Multilevel projection-based nested Krylov iteration for boundary value problems*, SIAM J. Sci. Comput., 30 (2008), pp. 1572-1595.
- [3] Y. A. ERLANGGA AND R. NABBEN, *Deflation and balancing preconditioners for Krylov subspace methods applied to nonsymmetric matrices*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 684-699.
- [4] E. EFSTATHIOU AND M. J. GANDER, *Why restricted additive Schwarz converges faster than additive Schwarz*, BIT Numerical Mathematics, 43 (2003), pp. 945-959.
- [5] L. GIRAUD AND S. GRATTON, *On the sensitivity of some spectral preconditioners*, SIAM J. Matrix Anal. Appl., 27 (2006), pp. 1089-1105.
- [6] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, 3rd Edition, John Hopkins University Press, Baltimore, MD, 1996.
- [7] P. GOSSELET AND C. REY, *On a selective reuse of Krylov subspaces in Newton-Krylov approaches for nonlinear elasticity*, In Domain decomposition methods in science and engineering, pages 419-426(electronic). Natl. Auton. Univ. Mex., México, 2003.
- [8] P. HAVÉ, R. MASSON, F. NATAF, M. SZYDLARSKI, H. XIANG, AND T. ZHAO, *Algebraic domain decomposition methods for highly heterogeneous problems*, to appear in SIAM J. Sci. Comput., 2013.
- [9] G. KARYPIS AND V. KUMAR, *A fast and highly quality multilevel scheme for partitioning irregular graphs*, SIAM J. Sci. Comput., 20 (1998), pp. 359-392.
- [10] R. B. MORGAN, *GMRES with deflated restarting*, SIAM J. Sci. Comput., 24 (2002), pp. 20-37.
- [11] F. NATAF, H. XIANG, V. DOLEAN, AND N. SPILLANE, *A coarse grid space construction based on local Dirichlet to Neumann maps*, SIAM J. Sci. Comput., 33 (2011), pp. 1623-1642.
- [12] R. NABBEN AND C. VUIK, *A comparison of deflation and the balancing preconditioner*, SIAM J. Sci. Comput., 27 (2006), pp. 1742-1759.
- [13] A. PADIY, O. AXELSSON, AND B. POLMAN, *Generalized augmented matrix preconditioning approach and its application to iterative solution of ill-conditioned algebraic systems*, SIAM J. Matrix Anal. Appl., 22 (2000), pp. 793-818.
- [14] M. L. PARKS, E. DE STURLER, G. MACKEY, D. D. JOHNSON, AND S. MAITI, *Recycling Krylov subspaces for sequences of linear systems*, SIAM J. Sci. Comput., 28 (2006), pp. 1651-1674.
- [15] O. PIRONNEAU, F. HECHT, A. LE HYRIC, AND J. MORICE, *FreeFem++*, Laboratoire J.-L. Lions, Université Pierre et Marie Curie, available online at <http://www.freefem.org/ff++/>, 3.7 edition, 2010.
- [16] Y. SAAD, *Iterative Methods for Sparse Linear Systems*, SIAM, 2003.
- [17] A. ST-CYR, M. J. GANDER, AND S. J. THOMAS, *Optimized multiplicative, additive, and restricted additive Schwarz preconditioning*, SIAM J. Sci. Comput., 29 (2007), pp. 2402-2425.

- [18] G. W. STEWART, *Matrix Algorithms Volume II: Eigensystems*, SIAM, 2001.
- [19] J. M. TANG, R. NABBEN, C. VUIK, AND Y. A. ERLANGGA, *Comparison of two-level preconditioners derived from deflation, domain decomposition and multigrid methods*, J.Sci.Comput., 39 (2009), pp. 340-370.